ELSEVIER

Biophysical Chemistry 96 (2002) 273–284

Biophysical Chemistry

www.elsevier.com/locate/bpc

A kinetic trap is an intrinsic feature in the folding pathway of single-chain Fv fragments☆

Wolfgang Hoyer[1], Kathrin Ramm, Andreas Plückthun*

Biochemisches Institut, Universität Zürich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland

Received 21 March 2001; received in revised form 25 October 2001; accepted 30 October 2001

Abstract

We have studied the equilibrium unfolding and the kinetics of folding and unfolding of an antibody scFv fragment devoid of *cis*-prolines. An anti-GCN4 scFv fragment carrying a V_L lambda domain, obtained by ribosome display, served as the model system together with an engineered destabilized mutant in V_H carrying the R66K exchange. Kinetic and equilibrium unfolding experiments indicate that the V_H mutation also affects V_L unfolding, possibly by partially destabilizing the interface provided by V_H, even though the mutation is distant from the interface. Upon folding of the scFv fragment, a kinetic trap is populated whose escape rate is much faster with the more stable V_H domain. The formation of the trap can be avoided if refolding is carried out stepwise, with V_H folding first. These results show that antibody scFv fragments do not fold by the much faster independent domain folding, but instead form a kinetically trapped off-pathway intermediate, which slows down folding under native conditions. This intermediate is characterized by premature interaction of the unfolded domains, and particularly involving unfolded V_H, independent of proline *cis–trans* isomerization in V_L. This work also implies that V_H should be a prime target in engineering well behaving antibody fragments.

Keywords: ScFv; Protein folding; Kinetic trap; Domain interface

1. Introduction

Considerable progress has been made in understanding and even predicting the folding kinetics of small monomeric proteins [1,2]. The folding of larger proteins consisting of multiple domains or subunits is, however, complicated by additional steps related to docking and annealing of domains which can be rate-limiting [3,4]. Antibodies are prototypes of multidomain proteins, since they are built from highly homologous building blocks, the

Abbreviations: ELISA, Enzyme-linked immunosorbent assay; GdmCl, Guanidine hydrochloride; ScFv, Single-chain antibody fragment consisting of the variable domains of the heavy and the light chain connected by a peptide linker; V_H, Variable domain of the heavy chain of an antibody; V_L, Variable domain of the light chain of an antibody.

☆ This paper is dedicated to Professor Rainer Jaeniche, a true scholar in the field of protein foliding.

*Corresponding author. Tel.: +41-1-635-5570; fax: +41-1-635-5712.

E-mail address: plueckthun@biocfebs.unizh.ch (A. Plückthun).

[1] These authors contributed equally to this study.

0301-4622/02/$ - see front matter
PII: S0301-4622(02)00022-4

immunoglobulin domains. One of the most attractive properties of this model system is the almost unlimited number of related molecules for testing the effects of structural variations of folding and stability. The smallest entity retaining the full antigen binding capacity is the Fv fragment consisting of the two variable domains of heavy and light chain, V_H and V_L. Because of the relatively weak affinity of V_H for V_L (K_D values between 10^{-6} M and 10^{-8} M) [5] they have usually been covalently linked by a peptide linker resulting in the single-chain Fv format (scFv). Besides their obvious biotechnological importance, antibody fragments represent an ideal model system for studying the folding behavior of multidomain proteins.

An additional advantage of the Fv and scFv antibody fragments as a model system to study domain interaction in protein folding lies in their favorable spectroscopic properties. On the one hand, both variable domains contain a conserved core tryptophan, whose fluorescence is almost completely quenched by the conserved disulfide-bond in the native fold. Refolding of the isolated domains thus results in a fluorescence decrease that can best be observed at 350 nm, the emission maximum of the unfolded state. On the other hand, the domains contain a variable number of additional tryptophans, which are usually located in the interface of the native heterodimer or even remain exposed in the antigen-binding site. In the isolated domains, all but the conserved core tryptophans are solvent exposed and thus do not contribute significantly to the fluorescence change observed upon refolding of the domain. Upon formation of the native heterodimer they become, however, buried in the hydrophobic interface and the fluorescence is accordingly blue-shifted, monitoring this change of environment. Formation of the native interface is thus associated with a large increase in fluorescence that is best observed approximately 320 nm [6]. These spectroscopic features facilitate the distinction and assignment of folding events greatly.

The process of structure formation during in vitro refolding of scFv and Fv fragments has been studied extensively in our laboratory over the past years using fragments derived from two different antibodies, the phosphorylcholine binding antibody McPC603 and the HER-2 binding antibody hu4D5-8 [6–14]. From these experiments a general folding scheme for such antibody fragments has emerged. Briefly, stable domains fold independently, followed by fast association to the native heterodimer. Less stable domains probably follow a similar pathway, but require the stabilizing influence of the large hydrophobic V_H/V_L interface for stable structure formation. A complication is introduced by the slow isomerization of the peptide bond preceding ProL95, which is close to the V_L/V_H interface and at least in the antibody McPC603 must be in *cis* for the native interface to form. The refolding of the scFv is slowed down in comparison to the refolding of the respective Fv fragment, due to premature interaction of the domains caused by the close proximity of the domains in the scFv format, leading to the population of a kinetic trap.

In the present study we have for the first time investigated a scFv which is devoid of *cis*-prolines in the native fold. We have used as model system a fragment obtained by ribosome display [15], which recognizes a peptide derived from the transcription factor GCN4 [16]. Its V_L domain is derived from a lambda type light chain, which never carry the almost universally conserved *cis*-prolines of kappa domains, L8 and L113, (according to a new consensus numbering scheme [17], corresponding to L95 according to the Kabat numbering scheme [18]). We now compare the folding kinetics in the absence of limiting proline *cis*/*trans* isomerization with the model for scFv folding to evaluate its general validity. Comparison of the folding kinetics of the wild type anti-GCN4 with a mutant destabilized in the V_H domain also allows some novel conclusions concerning the influence of the heterodimer interface on stability and the factors influencing the population of the kinetic trap during refolding.

2. Materials and methods

2.1. Protein expression and purification

The gene for the affinity matured anti-GCN4 scFv fragment with a 21-mer $G_5SG_4SG_4S_2G_3S$ linker in the orientation V_L-linker-V_H, obtained by

Table 1
Rate constants k (s^{-1}) for refolding and unfolding of the anti-GCN4 wt, the HR66K variant and the isolated V_L-domain

	final GdnHCl concentrations (M)				
	0.5	2.2	2.5	3	4
	Refolding		Unfolding		
anti-GCN4 wt	0.014 0.003	0.014	0.004 0.089	0.006 0.26	0.049
anti-GCN4 wt equilibrium intermediate	0.047				0.05
anti-GCN4 destabilized	0.005 0.0003	0.0015	0.008 0.15	0.006	0.05
anti-GCN4 dest. equilibrium intermediate	0.043				0.048
V_L-domain	0.04		0.061	0.17	0.29

ribosome display [15] was cloned between the *Sfi*I sites of the secretion vector pAK400 [19] for periplasmic expression with the *pelB* signal sequence. In the destabilized mutant ArgH66 (H66 according to Kabat [18], H77 according to the new consensus AHo numbering scheme [17]) was replaced by Lys and the resulting his-tagged gene was cloned into the secretion vector pIG6 [20]. The isolated domains were obtained by PCR of the respective wild-type domains and cloned between the *Sfi*I sites of the secretion vector pAK400.

All proteins were expressed in the periplasm of the *E. coli* strain JM83 at 25 °C. The isolated V_L domain (Mw 15235 Da) and both scFvs, the wt. (Mw 27663) and the R66K variant (Mw 27635), were purified from soluble cell extracts first with an immobilized metal ion affinity chromatography (IMAC) column at pH 7. The two scFvs were subsequently loaded onto an antigen-affinity column containing biotinylated GCN4(7P14P)-peptide attached to immobilized streptavidin and eluted at pH 2.2 [15]. The isolated V_L domain was purified by IMAC, followed by anion exchange chromatography at pH 8.0.

2.2. Analytical gel filtration

Analytical gel filtration was carried out on a Superdex-75 column on a SMART system in 50 mM Tris, pH 7.0, 150 mM NaCl with 0.005% Tween-20. The column was calibrated with bovine serum albumin, carbonic anhydrase and cytochrome c.

2.3. Fluorescence measurements

All fluorescence measurements were performed with a PTI Alpha Scan spectrofluorimeter (Photon Technologies Inc.) at 10 °C, using an excitation wavelength of 280 nm and emission wavelengths of 320 and 350 nm, as indicated. The buffer used was in all cases 50 mM Tris, pH 7.0, 150 mM NaCl (filtered, degassed), containing varying amounts of GdmCl. Protein concentrations were 0.2 μM for equilibrium measurements and 0.4 μM for kinetic measurements.

For equilibrium denaturation measurements, native protein was incubated overnight at 10 °C in Tris-buffer containing different amounts of GdmCl. Equilibrium transitions were followed by measuring intensities at 320 and 350 nm upon excitation at 280 nm, as well as the shift in emission maximum. Curves were not normalized, nor could ΔG values be derived, because of the obvious lack of two-state behavior.

For kinetic measurements the proteins, either native or unfolded as described in the text, were diluted rapidly (between 1:100 and 1:25) into Tris-

buffer containing GdmCl at various final concentrations. Kinetics were followed by the emission intensity at 320 or 350 nm upon excitation at 280 nm, as indicated. All kinetic traces were evaluated using the Kaleidagraph software (Synergy software, Reading, UK) with either a single exponential function or a sum of two exponential functions for biphasic kinetics.

3. Results

3.1. Protein purification, activity and spectroscopic properties

Both scFv fragments, the wild type (wt) and the variant carrying the destabilizing R66K mutation in the V_H domain (termed R66K), as well as the isolated V_L domain were purified from the soluble fraction of cell extracts. Both scFvs were entirely monomeric, as seen in gel filtration, and active, as determined by antigen-binding ELISA (data not shown). The V_H domain, on the other hand, had to be refolded from inclusion bodies and, since multimers were formed, it was not possible to isolate a monomeric species. Due to its additionally high tendency to aggregate, no measurements could be carried out with the isolated V_H domain.

It has previously been shown that the unfolding of a Fv or scFv fragment is associated with a red-shift in the emission maximum and a large decrease of the fluorescence intensity approximately 320 nm, which is primarily due to the destruction of the domain interface and the concomitant solvent exposure of the tryptophans buried in the native interface [6]. The unfolding of the isolated domains, on the other hand, leads to an increase of the fluorescence intensity at 350 nm, caused by the removal of the conserved tryptophan in the hydrophobic core of the domains from the quenching intradomain disulfide-bridge in the native structure [21]. In the case of the anti-GCN4 scFv, there are five tryptophans in addition to the single conserved core tryptophans of V_H and V_L. Two of these five are conserved residues located in the CDR2 of the V_H domain, which contribute to the domain interface and thus become solvent exposed upon its destruction. The other three are located in the CDR3s, two in the V_L and one in the V_H domain. Since these last three are solvent-accessible in the folded scFv, they should not contribute greatly to the fluorescence change observed upon folding. These spectroscopic properties thus provide an excellent means to distinguish domain folding from interface formation by the choice of the fluorescence emission wavelength.

3.2. Thermodynamic stability

GdmCl titrations of the two scFvs and the V_L domain were carried out by recording the shift in emission maximum, as well as the fluorescence intensity at 350 and 320 nm, to allow distinction between domain unfolding and destruction of the interface. The transition of both the wt and the R66K scFv is clearly not two-state. In the plot of the emission maxima (Fig. 1a), a small plateau is observed for the destabilized variant approximately at 2.2 M GdmCl, which seems to be shifted upwards to 2.6 M GdmCl in the case of the wt, where it is, however, less obvious. On the other hand, the two transitions can clearly be distinguished for both variants by measuring the fluorescence intensity. In Fig. 1b,c the fluorescence intensities are plotted at both wavelengths to facilitate the comparison. For both scFvs the first transition is associated with a large decrease of the intensity at 320 nm (Fig. 1b), indicative of the disruption of the native interface. In the case of the destabilized variant, this transition is shifted towards lower concentrations of GdmCl, suggesting a weakening effect of the mutation on the interface. The second transition leads to an intensity increase that can only be observed at 350 nm, indicating that this transition involves unfolding of an isolated domain. This second transition has different amplitudes for both variants, and the midpoint also appears to be shifted towards lower concentrations of denaturant in the case of the R66K mutant. Both transitions thus seem to be influenced by the R66K mutation in the V_H domain.

A comparison of the scFv transition with the unfolding transition of the isolated V_L domain (Fig. 1d) demonstrates that V_L unfolding occurs concomitantly with the destruction of the native

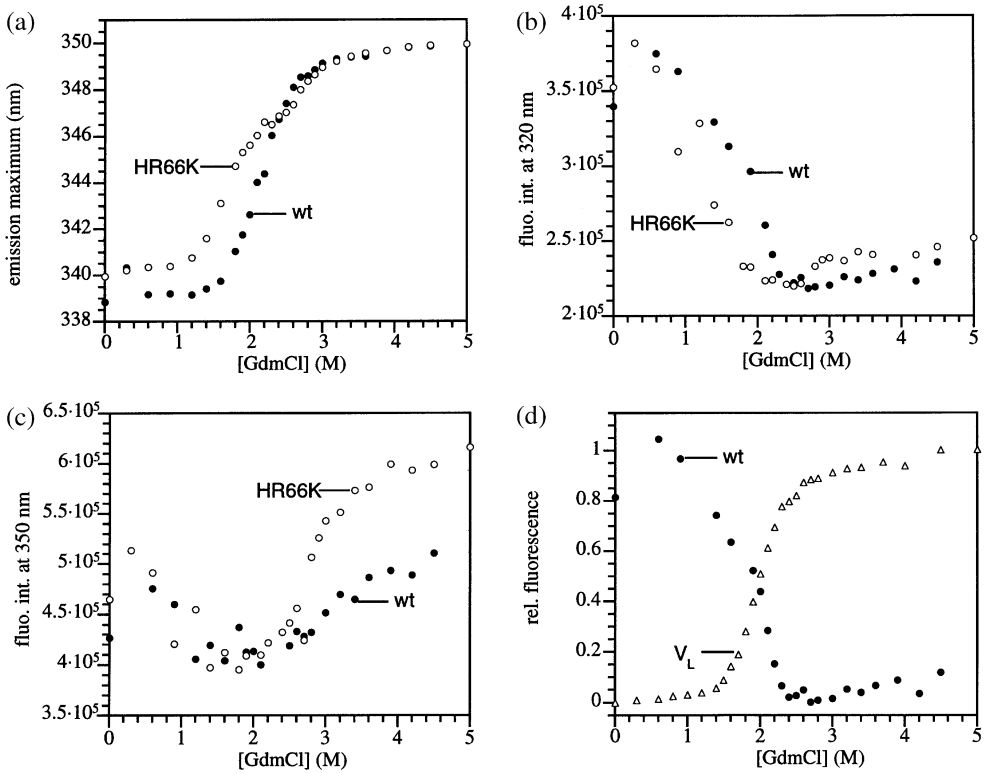

Fig. 1. GdmCl-induced equilibrium transitions at pH 7, 10 °C for both the wt anti-GCN4 (closed symbols) and the R66K mutant (open symbols). (a) Shift in emission maximum, (b) Fluorescence intensity at 320 nm, (c) Fluorescence intensity at 350 nm, (d) Comparison of the GdmCl-induced equilibrium transitions of the isolated V_L domain (△) at 350 nm and the wt anti-GCN4 at 320 nm (●).

interface, observed in the first transition of the wt scFv. Since the midpoint of the transition of the isolated wt V_L domain is approximately at 2 M GdmCl, it cannot be associated with the second transition that is observed for both scFvs at significantly higher GdmCl concentrations. Even though the V_H domain might stabilize V_L against unfolding, it could not do so if V_H unfolds before V_L. Therefore, the first transition must be assigned to V_L unfolding with concomitant loss of the interface. The second transition thus has to be attributed to the unfolding of the more stable V_H domain. As confirmed by kinetic measurements (see below), an equilibrium intermediate is therefore populated at 2.2 M GdmCl for both scFvs, which retains an at least partially structured V_H domain in the presence of an unfolded V_L domain.

Interestingly, destabilization of the V_H domain by the H-R66K mutation seems to influence both transitions observed in fluorescence (Fig. 1a–c). The stabilizing influence of Arg at position H66 had been identified from mutagenesis studies [22] in an antibody normally containing Lys. This residue is completely buried and does not make contact to the V_L domain. It is part of a conserved charge cluster within the V_H domain. Arg-H66 interacts strongly with Asp-H86, forming a buried salt-bridge. The destabilizing influence of the H-R66K mutation, located in the more stable V_H domain and remote from the interface, on the first transition is notable, as the native interface is destroyed concomitantly with the unfolding of the less stable V_L domain. The V_H core mutation must thus have long-range effects destabilizing the inter-

face and V_L. At the very least, the interface provided by the V_H domain of the less stable R66K mutant may not be completely intact at denaturant concentrations present in the transition region of V_L, thereby shifting the V_L transition to lower GdmCl concentrations, compared to the wt scFv. An implication is that the V_H interface must obviously be altered at intermediate GdmCl concentrations, meaning that the equilibrium intermediate, at least in the case of the destabilized variant, cannot consist of the combination of an entirely native V_H with an unfolded V_L.

3.3. Unfolding kinetics

The unfolding of both the wt- and the R66K mutant scFv proceeds in two distinct phases (Table 1). First, a fast phase with a large decrease of fluorescence intensity at 320 nm occurs, followed by a second much slower phase, characterized by a small fluorescence increase (Fig. 2a,b, circles). The second slow phase is not observed at denaturant concentrations below 2.5 M GdmCl for the wt and 2.2 M GdmCl for the R66K mutant (Fig. 2a and b, triangles). The sign and the amplitude of the fluorescence changes both indicate that the interface is destroyed in the faster phase, probably concomitant with the unfolding of the V_L domain, as suggested by the equilibrium transition, while the more stable V_H domain unfolds much more slowly and only at higher concentrations of denaturant. This view is supported by the unfolding kinetics of the isolated V_L domain at 2.5 M GdmCl, which occur with the same rate as the fast unfolding phase of the scFv at the same GdmCl concentrations (Fig. 2c). The phases have opposite sign, since in the case of the isolated V_L domain the dequenching of the core tryptophan is observed at 350 nm, while in the case of the scFv, the destruction of the interface, which goes along with V_L-unfolding, is dominant and most easily followed by fluorescence loss at 320 nm. The second slower phase of scFv unfolding taking place at higher denaturant concentrations, corresponding in its sign and amplitude to the unfolding of an isolated domain, must thus be due to the unfolding of the more stable V_H domain.

Unfolding of the R66K mutant corresponds exactly to the wt in sign and amplitude of all of the phases, supporting the existence of an equilibrium intermediate also for this variant (Fig. 2b). The rate of the slow phase, corresponding to the

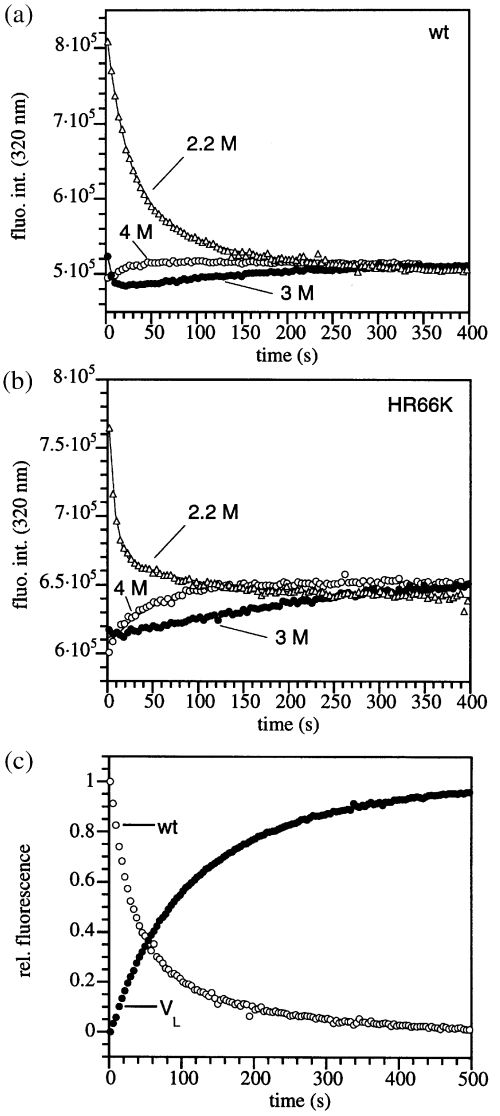

Fig. 2. Fluorescence traces of unfolding at 10 °C at 4 M (○), 3 M (●) and 2.2 M (△) GdmCl of (a) wt anti-GCN4 scFv, (b) R66K anti-GCN4 scFv. The intensity at 320 nm was followed. (c) Fluorescence traces of unfolding at 10 °C and 2.5 M GdmCl of wt anti-GCN4 scFv (○), and the isolated V_L domain (●). The intensity was followed at 320 nm (scFv) or 350 nm (V_L).

unfolding of the V_H domain, is identical for both variants, implying that the H-R66K mutation destabilizes the native and the transition state of the V_H domain by the same amount. Alternatively, the interaction of Arg66 could already be disrupted in the intermediate state, thus no more influencing the further denaturation of V_H. The only difference is seen in the rate of the first phase, leading to a large fluorescence decrease at 320 nm, which is faster for the H-R66K mutant. This is in accordance with a faster destruction of the interface between the two domains, which might be weakened by the mutation in V_H destabilizing the interface at the denaturant concentrations present in the experiments, as has been suggested from equilibrium measurements (see above).

3.4. Refolding kinetics are slow due to kinetic trapping

The scFv fragments we investigated so far [6–14] all had V_L domains of the κ-type containing two conserved *cis*-prolines in their native fold. Since the anti-GCN4 scFv has a λ-type V_L domain lacking *cis*-prolines in the native structure, its refolding should not be limited by proline isomerization. Nevertheless, refolding of the wt scFv from equilibrium is limited by a rather slow phase with a rate of approximately 0.003 s^{-1}. Double jump experiments, where refolding is started from short-term denatured protein retaining the prolines in their native conformation, confirmed that this slow phase is not due to proline *cis*/*trans* isomerization, since it yields exactly the same rate as after equilibrium denaturation (Fig. 3a). The fact that the rate of this slow phase is largely independent of the final denaturant concentration (data not shown) points towards the formation of a kinetic trap (see below). Therefore, the escape from this trap must be giving rise to this phase, rather than a true on-pathway folding phase. The formation of kinetically trapped intermediates has been observed for all scFv fragments investigated so far and could be attributed to premature interaction of the domains caused by the covalent linkage in the scFv format [12].

The refolding of the isolated V_L domain occurs on a much shorter timescale (Fig. 3b, Table 1),

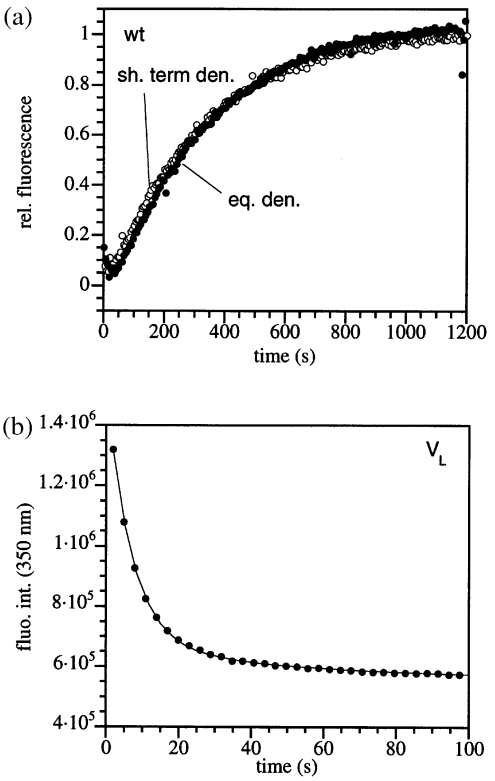

Fig. 3. (a) Fluorescence traces of refolding of wt anti-GCN4 scFv at 10 °C starting from equilibrium denatured protein (●) and short-term denatured protein (○). The intensity at 320 nm was followed. (b) Fluorescence trace of refolding of isolated V_L domain at 10 °C starting from equilibrium denatured protein. The intensity was followed at 350 nm.

confirming again the absence of *cis*-prolines in its native structure. V_L can thus not be responsible for the rate-limiting phase observed in the refolding of the anti-GCN4 scFv. Nevertheless, refolding of V_L from a trapped intermediate can contribute to the slow phase. When refolding of the scFvs is started from the equilibrium intermediate populated at 2.2 M GdmCl, refolding is fast and occurs with the same rate of approximately 0.05 s^{-1} for both the wt and the R66K variant (Fig. 4a). This corresponds well to the refolding rate of the isolated V_L domain (0.04 s^{-1}) (Fig. 3b) and supports the existence of an equilibrium intermediate with a largely intact V_H domain at approximately 2.2 M GdmCl, as indicated above. As discussed before for the example of unfolding in

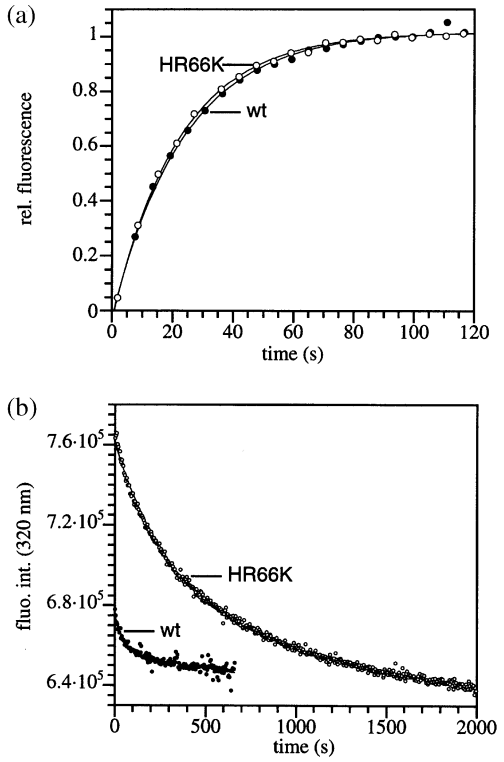

Fig. 4. Fluorescence traces of refolding at 10 °C of wt anti-GCN4 scFv (●) and the R66K variant (○) (a) in a final GdmCl concentration of 0.5 M, starting from the equilibrium intermediate at 2.2 M GdmCl, (b) in a final GdmCl concentration of 2.2 M starting from fully denatured protein. The intensity was followed at 320 nm.

Fig. 2c, the phases for refolding of the scFvs at 320 nm (Fig. 4a) and the isolated V_L domain at 350 nm (Fig. 3b) have opposite sign.

Unfortunately, the refolding of the isolated V_H domain could not be measured, due to its tendency to multimerize. From an analogy drawn with other V_H domains investigated so far [6], its refolding rate is most likely an order of magnitude faster than the rate-limiting phase of the wt scFv. Some support for this argument comes from refolding experiments of the scFvs carried out at 2.2 M GdmCl, a denaturant concentration where the V_L domain is not stable and thus only the V_H domain should be able to refold to a native-like structure. Under these conditions, a phase leading to a fluorescence decrease with a small amplitude at both 320 and 350 nm, as would be expected for the refolding of an isolated domain, is observed for the wt (Fig. 4b). The rate for the wt scFv (0.014 s^{-1}) even at this high concentration of denaturant is much faster than the slow phase of the wt scFv refolding under native conditions (0.003 s^{-1}) (Table 1). If the wt anti-GCN4 scFv is hence allowed to refold in a stepwise manner, with V_H refolding first at high concentrations of denaturant, followed by V_L and formation of the native interface under native conditions, the formation of a trapped off-pathway intermediate can be prevented. This confirms the earlier findings that trapping is due to premature domain interaction at low denaturant concentration, caused by the covalent linkage of the two domains in the scFv-format which increases their local concentration.

The refolding of the V_H domain of the R66K variant at intermediate GdmCl concentrations occurs with a 10-fold slower rate (0.0015 s^{-1}) than that of the wt anti-GCN4 (Fig. 4b), reflecting the reduced stability of the V_H domain of the R66K variant seen in the equilibrium transition. In the folding reaction, the H-R66K mutation thus influences the height of the transition state relative to the unfolded initial state, while in the unfolding reaction the height of the transition state with respect to the initial folded state is not changed. This implies that this residue makes an interaction that is crucial for the folding of the V_H domain.

Since the V_L domains of both scFvs are identical and since, moreover, they also refold with identical rates in the context of the different V_H domains in the wt and the R66K scFv (Fig. 4a), the large difference in the refolding rates of these two scFvs (Fig. 5a,b) has to be attributed to the contribution of the V_H domain. In accordance with the 10-fold slower rate of refolding of the destabilized V_H domain (Fig. 4b), the refolding of the R66K scFv under native conditions also occurs on a much longer timescale. Fig. 5a,b shows the refolding of the two scFvs from fully denatured protein in a final GdmCl concentration of 0.5 M, observed at two different wavelengths. At 350 nm a decrease of the fluorescence intensity is observed for both scFvs that occurs with a rate of 0.014 s^{-1} in the

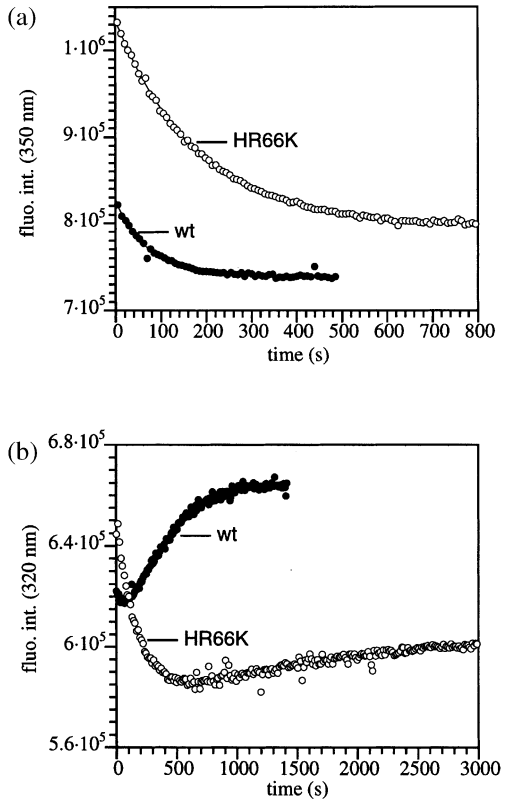

Fig. 5. Fluorescence traces of refolding at 10 °C of wt anti-GCN4 scFv (●) and the R66K variant (○) in a final GdmCl concentration of 0.5 M, starting from fully denatured protein. The intensity was followed in (a) at 350 nm and in (b) at 320 nm.

case of the wt and a rate of 0.005 s^{-1} in the case of the R66K variant. This phase is also seen at 320 nm (Fig. 5b), although with very different amplitudes for the two scFvs. At this wavelength, however, an additional much slower phase is observed with rates of 0.003 s^{-1} for the wt and 0.0003 s^{-1} for the R66K mutant (Table 1). Since this slow phase is only observed at 320 nm, it is likely that it involves the formation of the native interface. The mutation in the V_H domain thus appears to impair the recovery of the R66K scFv from the kinetic trap, suggesting that V_H folding is involved in the escape from the trap. A lower transition state of V_H folding depopulates the trapped intermediate more rapidly.

4. Discussion

The aim of this study was to complement our previous detailed investigations on the folding of various Fv and scFv fragments [6–14] and to test whether the proposed common folding pathway is also followed in the absence of limiting proline isomerization, which had previously complicated the analysis. For this purpose we have chosen a scFv fragment directed against a variant of the transcription factor GCN4 as a model system [15,16], since it contains a λ-type light chain devoid of *cis*-prolines in the native fold. Additionally, we have compared the folding behavior of the wt scFv with a variant that carries a destabilizing mutation in the V_H domain. Due to the unique influence of this mutation on both domains, we can now draw new conclusions relating to the off-pathway reactions during the refolding of the scFv fragment.

For both the wt anti-GCN4 scFv and the destabilized variant with the H-R66K mutation, an equilibrium intermediate is populated at approximately 2.2 M GdmCl in equilibrium unfolding experiments. It could be shown to consist of an unfolded V_L domain and a largely intact V_H domain. The fact that a destabilizing mutation in V_H, which is not located in the interface, is able to shift the first transition to lower denaturant concentrations is notable. Since this transition involves unfolding of the much less stable V_L and the concomitant loss of the interface, this implies that the interface provided by a destabilized V_H domain may partially 'melt' during this transition, thus destabilizing V_L even more. Caution in the use of the terms 'native' and 'unfolded' is therefore necessary with respect to the domains in the equilibrium intermediate. In the case of the destabilized R66K scFv at least, the V_H domain must also be affected to some extent in the equilibrium intermediate, since otherwise no influence on V_L-stability, mediated by the common interface, should be observed.

Although the folding of the isolated V_H domain could not be measured, it could be followed indirectly in the scFv format. The H-R66K mutation does not affect the unfolding rate of the V_H domain but slows the refolding rate by a factor of

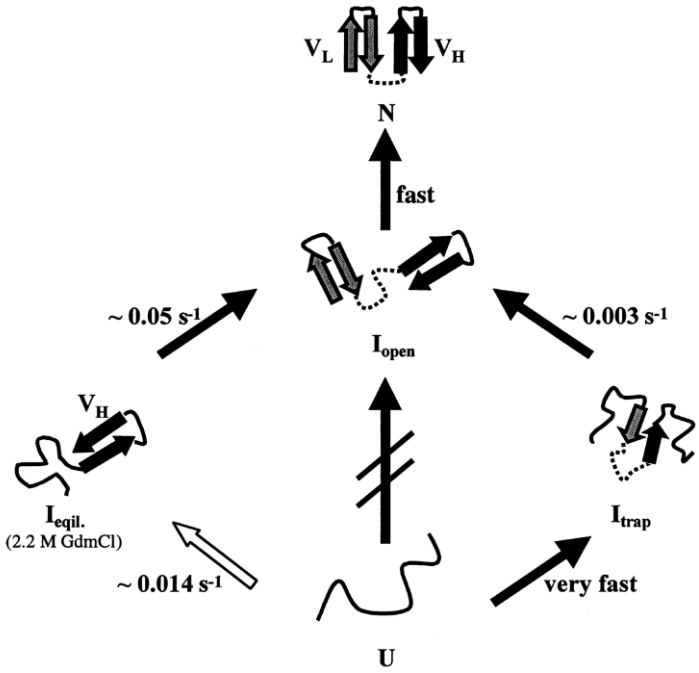

Fig. 6. Proposed refolding scheme for the wt anti-GCN4 scFv. Folded structure is symbolized by light-gray (V_L domain) and dark-gray arrows (V_H domain). The crossed-out arrow indicates that this direct step is not taking place. The thick black arrows denote reactions occurring under native conditions. The white arrow denotes a reaction which only occurs at intermediate denaturant concentrations (2.2 M GdmCl). For further details, see text.

ten, suggesting that the contact made by Arg66 in the wt is already formed in the transition state of folding of the V_H domain. Enforced sequential refolding, first of the V_H domain at high concentrations of denaturant where V_L is not stable, and subsequently of the V_L domain under native conditions accounts for all spectral changes observed upon refolding of the entire scFv from the fully unfolded state. This enforced sequential folding is faster than refolding of the entire scFv from the fully unfolded state under native conditions. It is thus clear that the slow phase, which is observed upon refolding of the fully denatured anti-GCN4 scFv at low concentrations of denaturant, does not stem from slow folding of the isolated domains, but is due to population of a kinetic trap under native conditions. The results thus confirm the hypothesis put forward earlier, stating that kinetic trapping is an intrinsic property of the scFv format, which is independent of other rate-limiting reactions and occurs due to premature domain interaction. Particularly, these results show that existence of this kinetic trap is independent of proline *cis–trans* isomerization reactions, as the current model system does not have any *cis*-prolines.

Fig. 6 depicts a summary of the general folding scheme for the anti-GCN4 wt scFv. Upon refolding from completely unfolded protein (black arrows), a trapped intermediate is formed in the dead time of manual mixing, which then slowly converts in a rate-limiting step to an open intermediate. The direct folding pathway U to I_{open} (crossed line in Fig. 6) not used in the scFv, but it is used, in Fv fragments [6,11], which fold much faster, and thus the linker enforces an 'unnecessary' intermediate I_{trap} off this main track. Likewise, population of the kinetic trap is prevented, when refolding of the anti-GCN4 scFv is started from an equilibrium intermediate (I_{equil}) retaining mostly native structure in V_H. We consider the trapped intermediate I_{trap} as 'off-pathway', although it is apparently on

the actually employed folding pathway of the scFv, since the fast, 'direct' pathway from U to I_{open} (crossed line in Fig. 6), with independent folding of V_H and V_L and subsequent association, is not used. This phenomenon is caused by the linker and may well be a frequent feature also in natural multidomain-proteins.

This kinetic behavior has essentially been observed for all scFvs, with the only difference that for some scFvs there seems to be an additional direct path from the unfolded state to the open intermediate, depending on the refolding conditions. Additional complexity is introduced in the case of those scFvs containing a V_L domain of the κ-type through *cis/trans* isomerization at the level of the unfolded state and in the trapped as well as in the open intermediate.

The interesting properties of the variant carrying the destabilizing mutation H-R66K allow additional new insights into the factors determining the formation and the rate of escape from the kinetic trap. In the destabilized mutant the rate-limiting folding phase for the scFv (I_{trap} to I_{open}) is slowed down by a factor of ten, compared to the wt, when refolding is started from fully denatured protein. The refolding rate of the destabilized V_H domain observed at intermediate concentrations of denaturant (U to I_{equil}) is also slowed down by a factor of ten compared to the wt V_H domain. Since refolding of the V_L domain is unaffected in the context of the wt and the mutant V_H domain, when refolding is started from the equilibrium intermediate (I_{equil} to I_{open}), it thus seems plausible that the refolding rate of the V_H domain directly affects the rate of escape from the kinetic trap.

In this context, it is interesting to note that the trap is not populated when refolding of either of the anti-GCN4 variants is started from the equilibrium intermediate I_{equil}, i.e. in the presence of an already largely folded V_H domain. In the case of the hu4D5-8 scFv in the same V_L-linker-V_H format, on the other hand, it could be shown that a trapped kinetic intermediate was populated when folding was started from intermediate GdmCl concentrations. In this case, however, an equilibrium intermediate exists, which retains native structure in V_L, but not in V_H. This might imply that formation of a trapped intermediate is more likely to be an intrinsic property of an unfolded V_H domain, which becomes manifest in the presence of high local concentrations of the V_L domain, as present in the context of the covalently linked scFv.

The stability and the refolding rate of the V_H domain might thus be the major factors in the scFv format determining the rate of rescue from the kinetically trapped intermediate. Since trapping is prevented upon refolding started from an intermediate with an at least partially structured V_H (I_{equil} to I_{open}), the V_H domain, moreover, seems to be critical for the initial formation of the kinetic trap. This is of course highly relevant for all applications of scFv fragments and has implications for further attempts to improve solubility and reduce the aggregation tendency of these important molecules, suggesting that V_H domains should be the major target of engineering efforts.

Acknowledgments

This work was supported by a grant from the Studienstiftung to W.H., a predoctoral grant from the Roche Research Foundation to K.R. and by the Schweizerische Nationalfonds grant 31-47302.96.

References

[1] K.W. Plaxco, K.T. Simons, I. Ruczinski, D. Baker, Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics, Biochemistry 39 (2000) 11177–11183.
[2] D. Baker, A surprising simplicity to protein folding, Nature 405 (2000) 39–42.
[3] R. Jaenicke, Folding and association of proteins, Prog. Biophys. Mol. Biol. 49 (1987) 117–237.
[4] R. Jaenicke, R. Seckler, Protein misassembly in vitro, Adv. Protein Chem. 50 (1997) 1–59.
[5] A. Plückthun, Mono- and bivalent antibody fragments produced in *Escherichia coli:* engineering, folding and antigen binding, Immunol. Rev. 130 (1992) 151–188.
[6] M. Jäger, A. Plückthun, Folding and assembly of an antibody Fv fragment, a heterodimer stabilized by antigen, J. Mol. Biol. 285 (1999) 2005–2019.
[7] C. Freund, A. Honegger, P. Hunziker, T.A. Holak, A. Plückthun, Folding nuclei of the scFv fragment of an antibody, Biochemistry 35 (1996) 8457–8464.
[8] C. Freund, P. Gehrig, T.A. Holak, A. Plückthun, Comparison of the amide proton exchange behavior of the rapidly formed folding intermediate and the native state

of an antibody scFv fragment, FEBS Lett. 407 (1997) 42–46.

[9] C. Freund, P. Gehrig, A. Baici, T.A. Holak, A. Plückthun, Parallel pathways in the folding of a short-term denatured scFv fragment of an antibody, Fold. Des. 3 (1998) 39–49.

[10] M. Jäger, A. Plückthun, The rate-limiting steps for the folding of an antibody scFv fragment, FEBS Lett. 418 (1997) 106–110.

[11] M. Jäger, A. Plückthun, Domain interactions in antibody Fv and scFv fragments: effects on unfolding kinetics and equilibria, FEBS Lett. 462 (1999) 307–312.

[12] M. Jäger, A. Plückthun, Direct evidence by H/D exchange and ESI-MS for transient unproductive domain interaction in the refolding of an antibody scFv fragment, Protein Sci. 9 (2000) 552–563.

[13] K. Ramm, P. Gehrig, A. Plückthun, Removal of the conserved disulfide bridges from the scFv fragment of an antibody: effects on folding kinetics and aggregation, J. Mol. Biol. 290 (1999) 535–546.

[14] M. Jäger, P. Gehrig, A. Plückthun, The scFv Fragment of the Antibody hu4D5-8: Evidence for Early Premature Domain Interaction in Refolding, J. Mol. Biol. 305 (2001) 1111–1129.

[15] J. Hanes, L. Jermutus, S. Weber-Bornhauser, H.R. Bosshard, A. Plückthun, Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries, Proc. Natl. Acad. Sci. USA 95 (1998) 14130–14135.

[16] C. Berger, S. Weber-Bornhauser, J. Eggenberger, J. Hanes, A. Plückthun, H.R. Bosshard, Antigen recognition by conformational selection, FEBS Lett. 450 (1999) 149–153.

[17] A. Honegger, A. Plückthun, Yet another numbering scheme for immunoglobulin variable domains: An automatic modeling and analysis tool, J. Mol. Biol 309 (2001) 657–670.

[18] Kabat, E.A., Wu, T.T., Perry, H.M., Gottesman, K.S., Foeller, C. (1991) in Sequences of Proteins of Immunological Interest (91-3242, N.P.N., Ed.), National Technical Information Services (NTIS).

[19] A. Krebber, S. Bornhauser, J. Burmester, et al., Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system, J. Immunol. Meth. 201 (1997) 35–55.

[20] L. Ge, A. Knappik, P. Pack, C. Freund, A. Plückthun, in: C.A.K. Borrebaeck (Ed.), Antibody engineering, Oxford University Press, 1995, pp. 229–266.

[21] Y. Goto, K. Hamaguchi, The role of the intrachain disulfide bond in the conformation and stability of the constant fragment of the immunoglobulin light chain, J. Biochem. 86 (1979) 1433–1441, (Tokyo).

[22] K. Proba, A. Wörn, A. Honegger, A. Plückthun, Antibody scFv fragments without disulfide bonds made by molecular evolution, J. Mol. Biol. 275 (1998) 245–253.